AD-1501 670

Copy 15 of 33 copies

AD-A267 253

IDA PAPER P-2734

A HANDBOOK OF COST RISK ANALYSIS METHODS

Philip M. Lurie, *Project Leader*

Matthew S. Goldberg
Mitchell S. Robinson

DTIC
ELECTE
JUL 27 1993
S B D

April 1993

Approved for public release; distribution unlimited.

93-16849

INSTITUTE FOR DEFENSE ANALYSES
1801 N. Beauregard Street, Alexandria, Virginia 22311-1772

IDA Log No. HQ 92-41686

DEFINITIONS

IDA publishes the following documents to report the results of its work.

Reports

Reports are the most authoritative and most carefully considered products IDA publishes. They normally embody results of major projects which (a) have a direct bearing on decisions affecting major programs, (b) address issues of significant concern to the Executive Branch, the Congress and/or the public, or (c) address issues that have significant economic implications. IDA Reports are reviewed by outside panels of experts to ensure their high quality and relevance to the problems studied, and they are released by the President of IDA.

Group Reports

Group Reports record the findings and results of IDA established working groups and panels composed of senior individuals addressing major issues which otherwise would be the subject of an IDA Report. IDA Group Reports are reviewed by the senior individuals responsible for the project and others as selected by IDA to ensure their high quality and relevance to the problems studied, and are released by the President of IDA.

Papers

Papers, also authoritative and carefully considered products of IDA, address studies that are narrower in scope than those covered in Reports. IDA Papers are reviewed to ensure that they meet the high standards expected of refereed papers in professional journals or formal Agency reports.

Documents

IDA Documents are used for the convenience of the sponsors or the analysts (a) to record substantive work done in quick reaction studies, (b) to record the proceedings of conferences and meetings, (c) to make available preliminary and tentative results of analyses, (d) to record data developed in the course of an investigation, or (e) to forward information that is essentially unanalyzed and unevaluated. The review of IDA Documents is suited to their content and intended use.

The work reported in this publication was conducted under IDA's Independent Research Program. Its publication does not imply endorsement by the Department of Defense, or any other Government agency, nor should the contents be construed as reflecting the official position of any Government agency.

UNCLASSIFIED

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
April 1993

3. REPORT TYPE AND DATES COVERED
Final Report, Oct 1991 – Apr 1993

4. TITLE AND SUBTITLE
A Handbook of Cost Risk Analysis Methods

5. FUNDING NUMBERS
IDA CRP 9000-720

6. AUTHOR(S)
Philip M. Lurie, Matthew S. Goldberg, Mitchell S. Robinson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311-1772

8. PERFORMING ORGANIZATION REPORT NUMBER
IDA Paper P-2734

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311-1772

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12A. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12B. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

When estimating costs, interval estimates are more informative than simple point estimates. Interval estimates are sometimes available using analytical methods. These methods may be either exact (for example, a sum of random variables) or approximate (for example, based on Taylor series). When analytical methods are not available, simulation may be used instead. This paper illustrates the application of both analytical and simulation methods to the three main types of cost analysis: cost-estimating relationships, work breakdown structures, and stochastic network models. In addition, the paper contains descriptions of some commercially-available software products that can be used to perform simulations for the types of cost-risk analysis described.

14. SUBJECT TERMS
Cost Analysis, Cost-Estimating Relationships, Work Breakdown Structure, Stochastic Models

15. NUMBER OF PAGES
50

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

UNCLASSIFIED

IDA PAPER P-2734

A HANDBOOK OF COST RISK ANALYSIS METHODS

Philip M. Lurie, *Project Leader*

Matthew S. Goldberg
Mitchell S. Robinson

April 1993

Approved for public release; distribution unlimited.

INSTITUTE FOR DEFENSE ANALYSES

IDA Independent Research Program

PREFACE

This paper was prepared by the Institute for Defense Analyses (IDA) under the IDA Independent Research Program. The objective of the task was to survey methods of cost-risk analysis for use by IDA and DoD researchers.

This work was reviewed within IDA by Bruce R. Harmon, Karen W. Tyson, and James Bui.

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

CONTENTS

I. Introduction 1

II. Cost Estimating Relationships 5

A. Analytical Method 6

B. Simulation Method 11

III. Work Breakdown Structures 13

A. Analytical Methods 14

1. Cost-Element Distributions Specified Directly 14

2. Cost-Element Distributions Determined From CERs 23

B. Simulation Methods 24

1. Cost-Element Distributions Specified Directly 24

2. Cost-Element Distributions Determined From CERs 25

IV. Stochastic Network Models 27

Appendix: Software For Running Simulations A-1

References Ref-1

Abbreviations Abbr-1

TABLES

1. First-Unit Cost WBS for 600-Pound UHF Satellite 15
2. Hypothetical Work Breakdown Structure — Example 1 20
3. Hypothetical Work Breakdown Structure — Example 2 21
4. Hypothetical Work Breakdown Structure — Example 3 21
5. Probabilities Associated With Network Activities and Events 30

FIGURES

1. Density Function of Total Cost 10
2. Cumulative Distribution Function of Total Cost 11
3. Analytical Versus Simulated Distributions of Total Cost 12
4. Density Function of Total Cost 16
5. Cumulative Distribution Function of Total Cost 17
6. Simulated and Fitted C.D.F.s for WBS in Table 2 22
7. Simulated and Fitted C.D.F.s for WBS in Table 3 22
8. Simulated and Fitted C.D.F.s for WBS in Table 4 23
9. Hypothetical Estimated and Actual Cost Risk Distributions 23
10. Example of Network for Development of Hypothetical Fighter Aircraft 28
11. Estimated Distribution of Total Cost From Network Simulation 31

I. INTRODUCTION

This handbook is intended as a guide for IDA analysts in performing cost risk analyses for the Department of Defense (DoD) and other government agencies. It is the culmination of many months of effort in reviewing the available literature on the subject and in talking to several experts in the field. It is not intended to be all-inclusive in the sense that it covers every conceivable type of cost risk analysis; rather, it emphasizes the most commonly used methods and addresses their advantages and shortcomings. Because this handbook is intended to serve as a general guide to performing cost risk analyses, the reader is cautioned against using it as a cookbook. To be thorough, a cost analysis must take account of specific characteristics and knowledge of the system being evaluated and must apply a certain measure of heuristic analysis.

All cost estimates for major defense acquisitions must now be accompanied by a formal risk analysis (DoD Instruction 5000.2, "Defense Acquisition Management Policies and Procedures"). Gone are the days when a single point estimate of cost would suffice. The reason for this is simple: point estimates of cost are almost always wrong. In the case of major defense programs, point estimates have frequently underestimated the true cost by a wide margin. The reasons for this include changes in the specifications or requirements as the program progresses, optimistic estimates of advances in technology, funding instability, and schedule slippage. Although a formal analysis of cost risk will not reduce the risk inherent in a program, it will help program managers understand the nature of the risks involved, and quantify and display the uncertainty associated with cost estimates. The result is a more realistic assessment of the funding required for a project and of the likelihood of exceeding the point estimate.

We have found the terms "cost risk" and "cost uncertainty" defined in many different ways in the literature. Risk is frequently defined as the occurrence of an outcome subject to a known pattern of random variation, i.e., the probability distribution from which the outcome is generated is known. Uncertainty, on the other hand, is often defined as the occurrence of an outcome subject to unknown random fluctuations. For the purpose of this report, we treat risk and uncertainty as interchangeable and give them the restricted definition of placing a probability distribution around a point estimate of total cost. We assume that the analyst has a method for producing a point estimate and needs a way to assess how uncertainties in various program features will manifest themselves in the distribution of total cost. This approach assumes that the analyst has correctly identified the factors that influence cost and taken them into account when producing the point estimate. If not, performing a cost risk analysis will be of little value

Although it may seem obvious, it is worth noting that the method for producing the point estimate must be the same as that for estimating the risk. In fact, the point estimate for total cost should be a by-product of the cost risk analysis. We mention this fact because the literature review identified situations where researchers first obtained a point estimate and then later applied a totally different method to assess the risk. For example, we observed situations where the most likely (i.e., the mode) cost was derived as the product of the most likely price and most likely quantity. Unless the distributions of both price and quantity are symmetric, the mode of the product is not the product of the modes. This is a pitfall that must be avoided when performing a cost risk analysis.

Our literature search identified both qualitative and quantitative methods for assessing risk. The qualitative methods (e.g., subjective assessments of low, medium, or high risk) are of most use when there is little or no historical data available or when firm requirements have not yet been established. They are most appropriate for assessing risk at the earliest stages of program conception when even subjective opinions are difficult to elicit. This report considers only quantitative methods where probability distributions on cost elements or drivers can be estimated from historical data or deduced from expert opinion. Quantitative methods can be either analytical or based on simulations. The former involves the mathematical determination of a total cost distribution from its component cost distributions. The latter involves the computer generation of random costs from component distributions and aggregation into a total cost distribution. The primary features of each method are as follows:

- ***Analytical Methods***:
  - If sufficient data are available, compute the first four moments (mean, variance, skewness, kurtosis) of each component cost and product moments (correlations and possibly higher-order cross-product terms) between component costs. If all component costs are independent, all product moments are zero.
  - If little or no data are available, component distributions must be derived from expert opinion. In this case, compute only the first two moments (because it is highly unlikely that any expert will be able to provide subjective estimates of skewness and kurtosis) and component cost correlations.

- Aggregate moments and fit a distribution (which distribution to use is discussed later) to the total cost.

- ***Simulation Methods***:
  - Generate many random selections from each component distribution. The number to generate will depend on the particular situation but usually at least 500 are recommended.
  - Aggregate resulting values to obtain the simulated distribution of total cost, summarized as either a histogram or cumulative distribution function, and associated descriptive statistics.

Note that we have not included prediction intervals as an alternative method for assessing risk. Although the literature is full of examples of risk being measured by prediction intervals, we do not consider them to be very useful statements of uncertainty. For example, stating there is a 95-percent chance that total cost will fall between one and five billion dollars is not very informative because the range of possible costs is so wide, and because no indication is given of the probabilities of exceeding specified costs. In our opinion, prediction intervals are too gross a summary measure of the distribution of total cost. They are not needed if the entire distribution can be estimated.

Analytical methods often have the disadvantage of being computationally complex, but once the computations are made, the analysis is done. There are many situations in practice, however, where the computations are so complex that they become intractable. Simulations, on the other hand, are easy to set up on a computer, but take more time (this can sometimes be a concern even on a high-speed computer if a large project is being simulated), and are not particularly well-suited to performing sensitivity analyses. Sensitivity analyses can take a great deal of time because a separate risk analysis encompassing many simulation iterations must be performed for each variation in the parameters being examined. Furthermore, it can be very difficult to generate correlated random variables except in the case of multivariate normality.

Although the literature search turned up many different ways for performing cost risk analyses, most can be categorized as direct applications or variations of the following three methods:

- cost-estimating relationships—use regression analysis to relate cost drivers to historical costs;
- work breakdown structures—apply probability distributions to cost elements and aggregate them into a total cost; and

- stochastic networks—break down the acquisition and management processes into their component activities and events in proper time sequence.

In the following chapters we describe these three methods in detail, give numerical examples of how to use them to perform cost risk analyses, and describe situations in which they might be applied. We present both the analytical and simulation approaches to each of these methods. Finally, the appendix describes some of the available software for running simulations.

II. COST-ESTIMATING RELATIONSHIPS

Cost-estimating relationships (CERs) involve the functional modeling of cost in terms of explanatory factors. The explanatory factors are usually proxies, such as weight and speed of an aircraft, for processes that directly determine cost but are difficult to measure in advance. For this reason, CERs must be regarded as correlational rather than causal models. All CERs are estimated by fitting a curve through historical data on costs and associated drivers. Frequently, however, historical data are insufficient to estimate a reliable CER. Furthermore, historical program data usually reflect the final program specifications and cost, and any schedule and technical difficulties encountered. The problem in accurately specifying a CER is that schedule and technical risk can rarely be sorted out (because of the unavailability of data) and the originally programmed cost and requirements are frequently unknown. For example, the costs for two aircraft with the same final specifications, but different initial requirements, are likely to be quite different because more effort must be expended if required technologies cannot be achieved or if requirements change. Thus, the primary pitfall in using a CER is model inadequacy.

Although it is important to recognize model inadequacy as a source of cost-estimating risk, its treatment is beyond the scope of this paper. We assume throughout this chapter that the analyst has accurately specified a CER that will enable the derivation of a point estimate of cost. The principal sources of risk in employing a CER are:

- model misspecification—schedule or technical risk not adequately accounted for, linear model used when a logarithmic form is more appropriate, cost drivers omitted or inadequately measured (e.g., aircraft maneuverability), and so on;
- extrapolation error—model no longer applies to current data, or model may not apply to entire range specified for independent variables;
- estimation uncertainty—model parameters and error variance are estimates, not actual values;
- uncertainty in cost drivers—cost drivers are subject to random fluctuations determined by a probability distribution; and
- assumption of independence of cost drivers.

A. ANALYTICAL METHOD

The analytical method for estimating a cost risk distribution involves a method called propagation of error [Tukey (1957), Seiler (1987)]. This method is based on the Taylor-Series expansion of a cost function about the means of the independent variables. The Taylor-Series expansion of a function in n dimensions can be written as:

$$f(x_1,\ldots,x_n) = f(\mu_1,\ldots,\mu_n) + \sum_{i=1}^{n}(f_i')(x_i - \mu_i) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(f_{ij}'')(x_i - \mu_i)(x_j - \mu_j) + \ldots \quad (1)$$

where

$$f_i' = f_i'(\mu_1,\ldots,\mu_n) = \left.\frac{\partial f(x_1,\ldots,x_n)}{\partial x_i}\right|_{\mu_1,\ldots,\mu_n},$$

$$f_{ij}'' = f_{ij}''(\mu_1,\ldots,\mu_n) = \left.\frac{\partial^2 f(x_1,\ldots,x_n)}{\partial x_i \partial x_j}\right|_{\mu_1,\ldots,\mu_n},$$

and $\mu_i = E(x_i)$, $i = 1, 2,\ldots, n$.

Taking expectations of both sides of equation (1) gives:

$$E[f(x_1,\ldots,x_n)] \approx f(\mu_1,\ldots,\mu_n) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(f_{ij}'')Corr(x_i,x_j)\sqrt{V(x_i)V(x_j)} \quad (2)$$

Furthermore, if we drop the second-order terms in equation (1), square both sides, and take expectations, we obtain:

$$V[f(x_1,\ldots,x_n)] \approx \sum_{i=1}^{n}(f_i')^2 V(x_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(f_i')(f_j')Corr(x_i,x_j)\sqrt{V(x_i)V(x_j)} \quad (3)$$

In equations (2) and (3), the means and variances of the independent variables are those for the system being costed, not those of the sample data from which the CER was derived. When a new system is being costed, the values of the independent variables are rarely known with certainty. Both the means and variances of the independent variables are consequently derived from expert opinion. The correlations, on the other hand, are more likely to be estimated from historical data.

Note that if $f(x_1,\ldots,x_n) = x_1 + \cdots + x_n$, equations (2) and (3) reduce to the standard equations for the mean and variance of a sum. Although more accurate approximations for the mean and variance can be obtained by considering higher-order terms, the expressions will involve product moments (i.e., moments of the form $E(x_i^p x_j^q)$ for positive integers p and q, with $p + q > 2$) of the cost drivers [Seiler (1987)] that cannot be reliably estimated with small amounts of data. If the analyst is fortunate enough to have a large number of observations at his disposal, it would be wise to consider the higher-order terms to improve the estimates of the mean and variance.

We illustrate this procedure with an actual CER[1] for the first-unit cost (in thousands of dollars) of an unmanned spacecraft:

$$Cost = 5.48(SC\_WT)^{.94}(BOLP)^{.30}\,\varepsilon\,,$$

where SC_WT is the spacecraft dry weight (in pounds) and $BOLP$ is the beginning-of-life power (in watts). From the historical data base used to estimate the CER, the correlation between weight and power was estimated to be .77. The CER was estimated by taking logarithms of both sides and performing an ordinary least-squares regression. On the logarithmic scale, the mean square error of the regression was estimated to be $\sigma^2 = .13$. If ε is assumed to be log-normally distributed with $E(\log \varepsilon) = 0$, the mean and variance of ε are:

$$E(\varepsilon) = e^{\sigma^2/2}$$

and

$$V(\varepsilon) = e^{\sigma^2}(e^{\sigma^2} - 1)\,.$$

From the above equations, the mean and variance of ε are 1.07 and .16, respectively.

Now suppose we want to use this CER to estimate the cost of a new surveillance satellite for tracking ballistic missile launchers. Suppose further that the following information is elicited for the parameters of the new system:[2]

Variable	Low	Mode	High	Mean	Variance
Weight	5,000	6,500	7,000	6,167	180,556
Power	1,800	2,000	2,500	2,100	21,667

Finite bounds on ε can be obtained by adding and subtracting three standard deviations to the mean on the logarithmic scale (i.e., zero), and transforming back to the original scale. The resulting bounds are $.33 < \varepsilon < 2.99$.

The spacecraft CER that was estimated is of the form

$$Cost = \alpha x_1^{\beta_1} x_2^{\beta_2}\,\varepsilon\,. \tag{4}$$

We therefore compute the following quantities:

---

[1] The data used to estimate this CER were obtained from the Air Force Unmanned Space Vehicle Cost Model (USCM) data base.

[2] In actuality, it is more common to elicit information on only the low, high, and modal values for each variable. From these values, we computed the mean and variance, assuming a Triangular distribution (discussed in the next chapter), because they are needed for the propagation-of-error formulas.

$$\frac{\partial Cost}{\partial x_1} = \alpha\beta_1\mu_1^{\beta_1 - 1}\mu_2^{\beta_2}E(\varepsilon) = (5.48)(.94)(6167^{-.06})(2100^{.30})(1.07) = 32.40 ,$$

$$\frac{\partial Cost}{\partial x_2} = \alpha\beta_2\mu_1^{\beta_1}\mu_2^{\beta_2 - 1}E(\varepsilon) = (5.48)(.30)(6167^{.94})(2100^{-.70})(1.07) = 30.37 ,$$

$$\frac{\partial Cost}{\partial \varepsilon} = \alpha\mu_1^{\beta_1}\mu_2^{\beta_2} = (5.48)(6167^{.94})(2100^{.30}) = 198,662 ,$$

$$\frac{\partial^2 Cost}{\partial x_1^2} = \alpha\beta_1(\beta_1 - 1)\mu_1^{\beta_1 - 2}\mu_2^{\beta_2}E(\varepsilon) = (5.48)(.94)(-.06)(6167^{-1.06})(2100^{.30})(1.07) = -.00032 ,$$

$$\frac{\partial^2 Cost}{\partial x_2^2} = \alpha\beta_2(\beta_2 - 1)\mu_1^{\beta_1}\mu_2^{\beta_2 - 2}E(\varepsilon) = (5.48)(.30)(-.70)(6167^{.94})(2100^{-1.70})(1.07) = -.01 ,$$

$$\frac{\partial^2 Cost}{\partial \varepsilon^2} = 0 ,$$

$$\frac{\partial Cost}{\partial x_1 \partial x_2} = \alpha\beta_1\beta_2\mu_1^{\beta_1 - 1}\mu_2^{\beta_2 - 1}E(\varepsilon) = (5.48)(.94)(.30)(6167^{-.06})(2100^{-.70})(1.07) = .00463 ,$$

$$\frac{\partial Cost}{\partial x_1 \partial \varepsilon} = \alpha\beta_1\mu_1^{\beta_1 - 1}\mu_2^{\beta_2} = (5.48)(.94)(6167^{-.06})(2100^{.30}) = 30.28 ,$$

and

$$\frac{\partial Cost}{\partial x_2 \partial \varepsilon} = \alpha\beta_2\mu_1^{\beta_1}\mu_2^{\beta_2 - 1} = (5.48)(.30)(6167^{.94})(2100^{-.70}) = 28.38 .$$

From equation (2), the mean cost is

$$E(Cost) = (5.48)(6167^{.94})(2100^{.30})(1.07) + (.5)[(-.00032)(180556) + (-.01)(21667)]$$
$$+ (.00463)(.77)\sqrt{(180556)(21667)} = 212,655 ,$$

and from equation (3) the variance is

$$V(Cost) = (32.40)^2(180556) + (30.37)^2(21667) + (198662)^2(.16)$$
$$+ (2)(32.40)(30.37)(.77)\sqrt{(180556)(21667)}$$
$$= 6,618,958,997 \text{ (standard deviation} = 81,357) .$$

From the lower and upper bounds on x_1, x_2, and ε, we obtain the lower and upper bounds on total cost as

$$Cost_l = (5.48)(5000^{.94})(1800^{.30})(.33) = 51,393 ,$$

and

$$Cost_u = (5.48)(7000^{.94})(2500^{.30})(2.99) = 705,060 .$$

We now have all the information we need to fit a distribution to the total cost. To be applicable to a wide variety of cost-estimating situations, a good candidate distribution left-skewed, or symmetric depending on its parameter values. Furthermore, because finite bounds are usually placed on the total cost, the distribution should be defined over a finite domain. The most commonly used distribution with these properties is the Beta distribution. The Beta distribution is specified as:

$$f(y)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)(h-l)^{\alpha+\beta-1}}(y-l)^{\alpha-1}(h-y)^{\beta-1}\ ,\ l\le y\le h\ ,$$

where $\alpha>0$ and $\beta>0$ are shape parameters, l is the lower bound, and h is the upper bound of the distribution. Because few analysts can relate to and specify the parameters of a Beta distribution, the parameters are usually estimated by the Method of Moments. The first step in this process is for the analyst to specify the endpoints of the distribution (the lowest and highest possible costs), either the modal or mean cost, and the variance of cost. The mode, mean, and variance of the Beta distribution are, respectively:

$$m=l+(h-l)\frac{1-\alpha}{2-\alpha-\beta}, \tag{5}$$

$$\mu=l+(h-l)\frac{\alpha}{\alpha+\beta}, \tag{6}$$

and

$$\sigma^2=(h-l)^2\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \tag{7}$$

Once the mean and variance have been specified, we set equations (6) and (7) to the numerical values of these quantities. This results in the following formulas for α and β in terms of the specified values for μ and σ^2 :

$$\alpha=\frac{(\mu-l)^2(h-\mu)}{\sigma^2(h-l)}-\frac{\mu-l}{h-l}, \tag{8}$$

$$\beta=\frac{\alpha(h-\mu)}{(\mu-l)}. \tag{9}$$

If the mode and variance are specified instead, the formulas become much more complex and are usually estimated by numerical means, for example, by using Microsoft *Excel*'s "Solver" routine.

Applying the Method of Moments to the example above, the estimated Beta parameters are $\alpha=2.71$ and $\beta=8.28$. The resulting density function and cumulative distribution function (c.d.f.) of total cost are shown in Figures 1 and 2, respectively. The mode of the total cost density shown in Figure 1 is 175,728. Note that this is *not* the cost obtained by substituting the most likely values for weight and power into the regression. The latter cost is:

$$\begin{aligned} Cost &= 5.48(6500^{94})(2000^{30})(1.07) \\ &= 220{,}096\ . \end{aligned}$$

Some analysts wrongly refer to this point estimate as the "most likely" cost, which refers to the mode of the distribution. As we have shown, however, the most likely cost in this example is 175,728. The reason it is smaller than the value obtained by substituting the most likely values for weight and power into the regression is that the distribution of weight, which dominates the regression, is skewed to the left. Therefore, it is necessary to perform a cost risk analysis even if all that is wanted is the most likely estimate of total cost.

Although the most likely estimate of total cost is 175,728, Figure 2 indicates that the probability of exceeding this cost is about 63 percent. This example vividly illustrates the pitfalls of relying on a point estimate alone. When a budgeting target is needed, an alternative to selecting the most likely value is to select the cost that gives an acceptable level of risk. For example, if 25 percent is considered to be an acceptable level of risk, the point estimate that gives this is 263,835—50 percent higher than the most likely estimate.

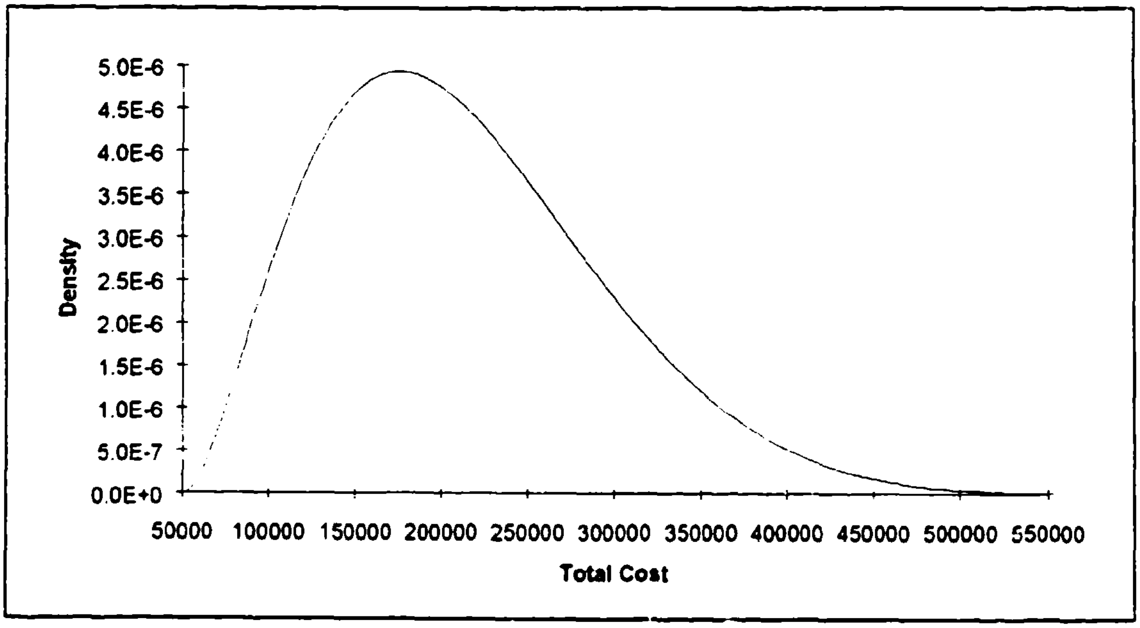

Figure 1. Density Function of Total Cost

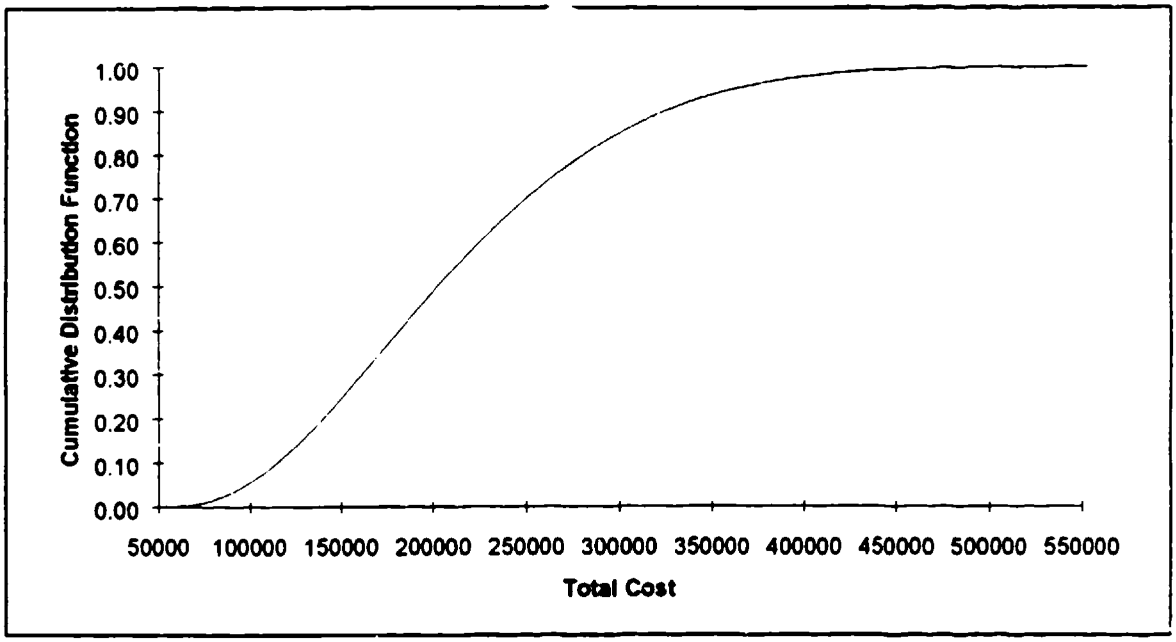

Figure 2. Cumulative Distribution Function of Total Cost

B. SIMULATION METHOD

Whereas the propagation of error method uses the first two moments[3] of each of the cost drivers and the error term, a simulation involves specifying distributions for these quantities. Commonly, Beta or Triangular distributions are used for the cost drivers and a log-normal distribution (if the CER is in logarithmic form) is used for the error term. The information embodied in a probability distribution is more complete than summary measures such as the first two moments. One might then expect that performing a simulation would result in a more accurate assessment of the distribution of total cost. This presupposes, however, that the distributions placed on the cost drivers and the error term are correct. In all likelihood, they are not (they are merely convenient ways of expressing subjective knowledge), and there is no guarantee that the resulting cost distribution will be any more accurate.

To illustrate this procedure, consider the same example used in the section on propagation of error. Equating the moments specified for weight and power to the theoretical moments of a Beta distribution allows us to determine the required distribution parameters. The cost drivers—weight and power—are then generated from the resultant Beta distributions and the error term is generated from a log-normal distribution. One

[3] Third- and fourth-order moments can be used as well if sufficient data are available. However, the formulas are much more complex and a four-parameter distribution, unlike the two-parameter Beta, must be used to fit the moments of total cost.

thousand iterations of this procedure were performed with the total cost determined from equation (4) at each iteration. The empirical c.d.f. was calculated as a step function with increments of 1/1000 at each of the ordered simulated costs. The result is superimposed over the analytically-derived c.d.f. and shown in Figure 3. As can be seen, the two curves are very similar; that is, the analytical and simulation techniques give essentially the same result. This need not be true in general, however.

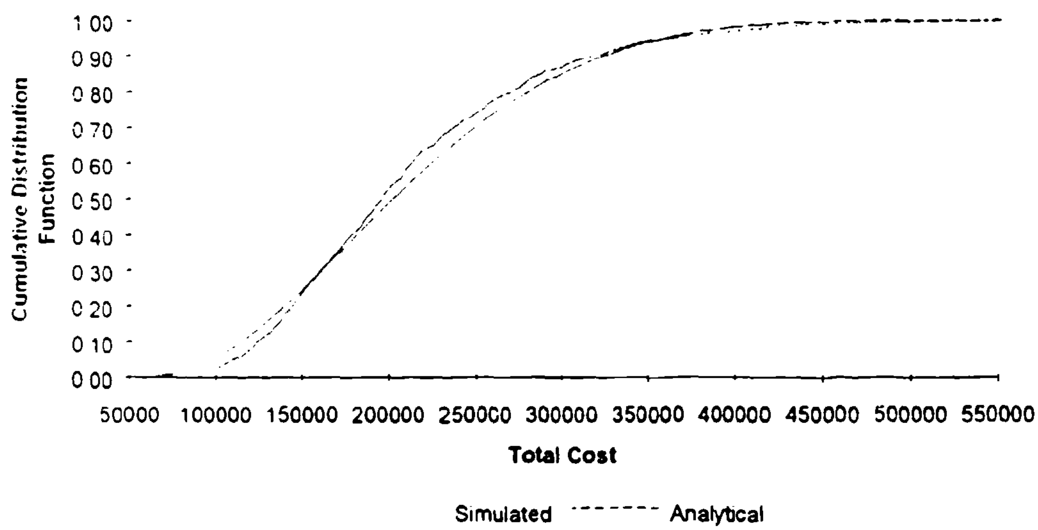

Figure 3. Analytical Versus Simulated Distributions of Total Cost

The most difficult aspect of simulating a CER is reproducing the correlation structure and moments of the cost drivers while maintaining their bounds (i.e., the low and high values for each variable). Unfortunately, fully tested and reliable methods for accomplishing this are not yet available.[4] The cost drivers in the above example were generated using a new methodology proposed in Lurie and Goldberg (1992), the details of which are beyond the scope of this paper. The methodology looks promising but requires more research before we would feel confident enough to recommend it. In the absence of any other acceptable method for simulating correlated random variables, we recommend using the propagation of error method, tedious though it may be, to estimate the total system cost risk.

[4] The method of Li and Hammond (1975) has been available for some time, but it requires extremely complex computations. Their method is probably impractical for simulating correlated Beta distributions.

III. WORK BREAKDOWN STRUCTURES

A work breakdown structure (WBS) is a hierarchical system of subordinate-level cost elements that are directly related to activities that define a project under development or production. As such, it completely defines the costs associated with the system under consideration and is therefore a causal model of total cost. Whereas CERs are relatively easy to construct and estimate, a WBS requires data at much greater levels of detail. A WBS also takes more time to construct and may require the inputs of numerous subject-matter experts, each knowledgeable in a particular aspect of project operation. Frequently, the "data" take the form of expert opinion. If expert opinion is accurate at lower levels of indenture, a WBS can potentially provide a more realistic assessment of cost risk than a CER. There is the danger, however, of disaggregating the data at too great a level of detail. Consequently, errors are compounded at higher levels of detail and may add to the level of uncertainty.

The primary sources of risk for a WBS are:

- disaggregating too much (using too low a level of indenture for accurate cost distributions to be provided),
- assuming independence of cost elements,
- assuming independence of price and quantity (for example, schedule slippage may require more manhours at overtime wages, resulting in a positive correlation between number of manhours and cost per manhour), and
- ignoring external cost drivers (external influences, such as funding instability, shifting requirements, and poor subcontractor performance, need to be taken into account).

To address the last point, analysts frequently use CERs to model the lower-level cost elements in a WBS.

The analysis of risk in a WBS begins with the specification of probability distributions on lower-level cost elements. Sometimes the lower-level cost distributions are formed as the product of separate distributions on price and quantity. The most commonly used distributions for specifying cost risk are the Beta and Triangular distributions. If a Beta distribution is specified, its parameters may be estimated by the Method of Moments as shown in equations (8) and (9) in the previous chapter. The Triangular distribution is used perhaps even more frequently than the Beta distribution because the only required parameters are the mode and the endpoints. These parameters directly determine the shape of the distribution and no indirect estimation of parameters is necessary. The mean and variance of the Triangular distribution are:

$$\mu = \frac{l+m+h}{3} , \tag{10}$$

and

$$\sigma^2 = \frac{(h-l)^2 + (m-h)(m-l)}{18} . \tag{11}$$

The primary advantage of the Triangular distribution—its simplicity—is also its primary disadvantage because its flexibility is limited. Unlike the Beta distribution, the Triangular distribution has only three parameters and its variance is predetermined once these parameters are specified. As a consequence, the Triangular distribution generally has fatter tails and this implies an increased likelihood of observing costs near the endpoints.

A. ANALYTICAL METHODS

1. Cost-Element Distributions Specified Directly

When cost-element distributions are specified directly, that is, without the aid of a CER, the method for estimating the total-cost distribution is straightforward. The first two moments of the cost-element distributions are calculated, aggregated, and fit with a Beta distribution. The Beta distribution is used for its flexibility, not because we have any *a priori* notion that this is the actual total-cost distribution.

To illustrate this procedure, consider an aggregate Work Breakdown Structure for the first-unit cost (in thousands of dollars) of a 600-pound ultra-high frequency (UHF) satellite, consisting of 10 elements.[5] We placed a Triangular distribution on each of these elements as shown in Table 1.

[5] The data for this WBS were obtained from the Air Force Unmanned Space Vehicle Cost Model (USCM) data base.

Table 1. First-Unit Cost WBS for 600-Pound UHF Satellite

	Cost (Thousands of Dollars)				
Cost Element	Lower Bound	Mode	Upper Bound	Mean	Standard Deviation
Attitude Control	1,676	1,942	2,453	2,024	203
Electrical Power Supply	3,469	4,029	5,287	4,264	472
Telemetry, Tracking and Command	860	986	1,671	1,172	203
Structure and Thermal	366	576	963	635	156
Apogee Kick Motor	201	287	402	297	53
Digital Electronics	5,433	6,791	8,828	7,017	891
Communications Payload	2,228	2,475	3,713	2,805	374
Integration and Assembly	544	691	1,011	749	121
Program Support	10,410	12,428	17,400	13,413	1,809
Launch Operations and Orbital Support	639	792	1,030	820	103

The means and standard deviations in Table 1 were computed from equations (10) and (11). We also calculated correlations among the cost elements from historical data.[6] The correlation matrix is shown below.

	1	2	3	4	5	6	7	8	9	10
1	1.00	0.47	0.36	0.76	0.10	0.13	0.58	0.37	0.81	0.54
2	0.47	1.00	0.37	0.79	0.33	0.43	0.55	0.37	0.52	0.22
3	0.36	0.37	1.00	0.58	0.52	0.68	0.22	-0.05	0.26	-0.06
4	0.76	0.79	0.58	1.00	0.23	0.38	0.55	0.41	0.70	0.33
5	0.10	0.33	0.52	0.23	1.00	0.82	0.64	0.25	0.25	0.06
6	0.13	0.43	0.68	0.38	0.82	1.00	0.29	0.05	0.03	0.06
7	0.58	0.55	0.22	0.55	0.64	0.29	1.00	0.62	0.85	0.34
8	0.37	0.37	-0.05	0.41	0.25	0.05	0.62	1.00	0.58	0.64
9	0.81	0.52	0.26	0.70	0.25	0.03	0.85	0.58	1.00	0.39
10	0.54	0.22	-0.06	0.33	0.06	0.06	0.34	0.64	0.39	1.00

Using the standard formulas for the mean and variance of a sum,

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) \ ,$$

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} Corr(X_i, X_j)\sqrt{V(X_i)V(X_j)} \ ,^{7}$$

[6] Caution must be exercised when estimating a correlation matrix from historical data if the pairwise correlations are not all based on the same observations. The result may be an inconsistent correlation matrix. See Lurie and Goldberg (1992) for a further discussion of this topic.

[7] We have come across statements in seminars and in the literature to the effect that this formula is an approximation, or that it applies only in the case of the normal distribution. This is not true. The formula is *exact* and requires only that the first two moments of each component distribution exist.

we obtained the mean and standard deviation of total cost (in thousands of dollars) as 33,194 and 3,230, respectively. If we had mistakenly assumed independence among the cost elements, the calculated standard deviation would have been only 2,136. Also, the lower bound on total cost is 25,626 and the upper bound is 42,758. From equations (8) and (9), the estimated Beta distribution parameters are $\alpha = 2.50$ and $\beta = 3.25$ (if we had assumed independence, the estimated parameters would have been $\alpha = 6.29$ and $\beta = 8.16$). The resulting density function and c.d.f. are shown in Figures 4 and 5, respectively.

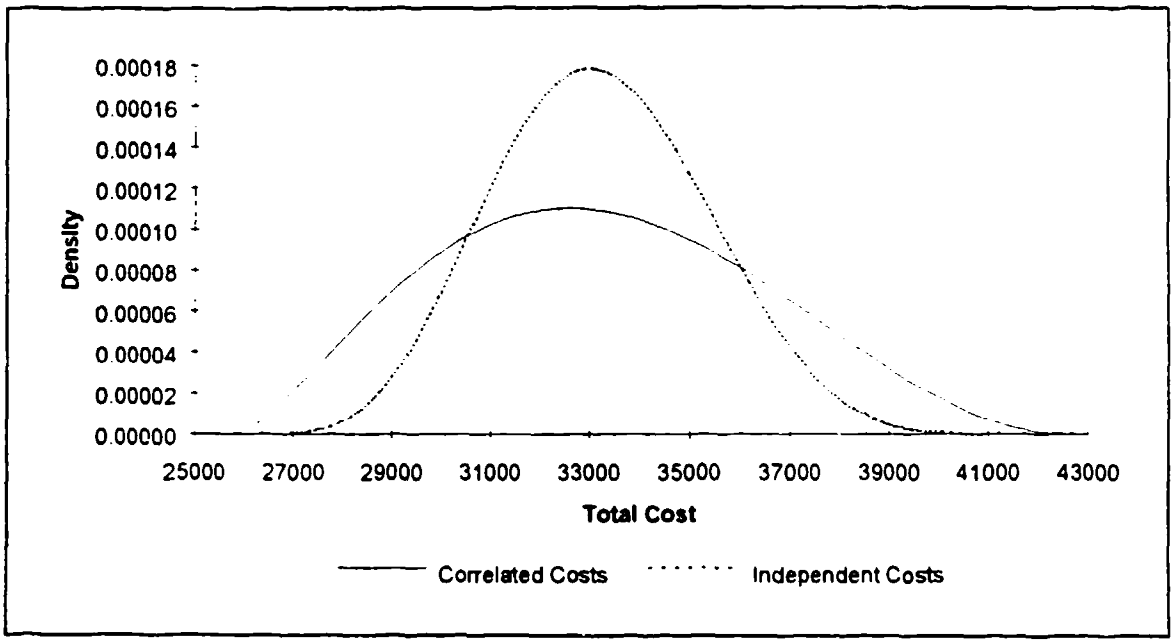

Figure 4. Density Function of Total Cost

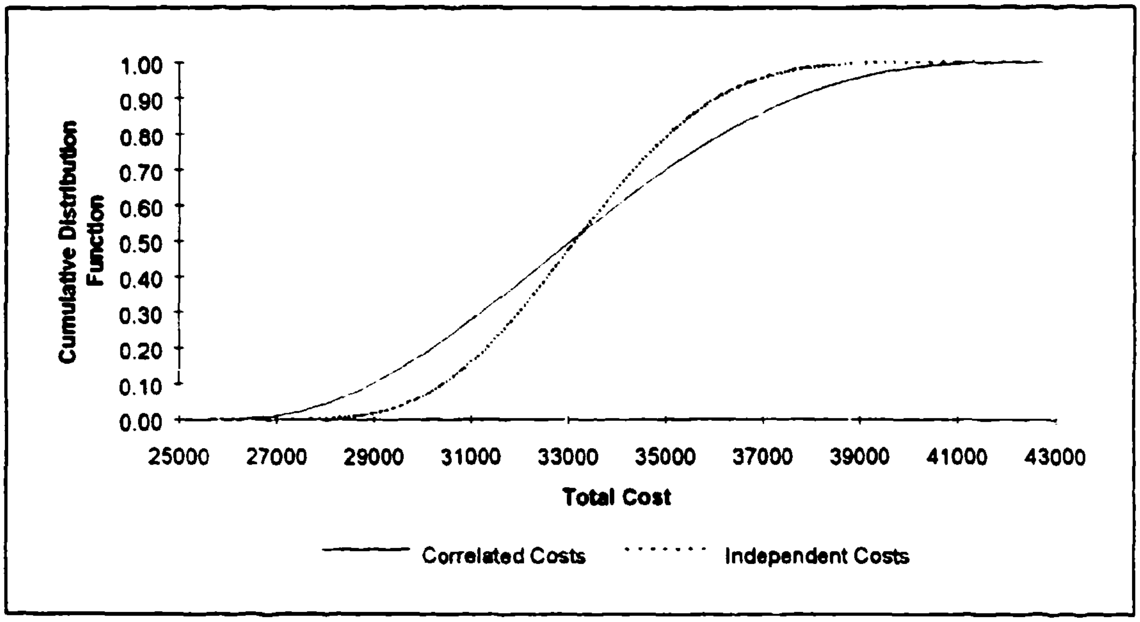

Figure 5. Cumulative Distribution Function of Total Cost

Figures 4 and 5 clearly demonstrate the consequences of wrongly assuming independence. In this example, the result is to greatly underestimate the probability of exceeding costs greater than the mean. Also note the near symmetry of the total-cost distribution, despite the fact that all ten component distributions are skewed to the right. Furthermore, the Beta distribution with parameter values both greater than or equal to 2 is very well-approximated by a normal distribution [Johnson and Kotz (1970)].

This calls to mind the Central Limit Theorem (CLT), which states that the sum of independent random variables with finite means and variances is asymptotically normally distributed, provided that the total variance is not dominated by only a few variables (see Feller (1971) for a more rigorous statement of the CLT). There is also a multivariate CLT that implies that the total-cost distribution is approximately normal when certain types of dependencies exist (e.g., when each cost element is a sum of independent sub-elements, but cost elements may have at least one sub-element in common). We may also expect the CLT to hold if the sum of the correlated cost elements is small relative to the total cost. It is of interest to observe that the nearly-symmetric cost distribution shown in Figure 4 is composed of only ten cost elements. Although the conventional wisdom is that most total-cost distributions should be skewed to the right, this example suggests this may not necessarily be the case.

To examine this further, consider three hypothetical Work Breakdown Structures, shown in Tables 2 through 4. Each WBS consists of costs decomposed as unit price ×

quantity. All price and quantity distributions are specified as independent right-skewed Triangular distributions.[8] The product of price and quantity is therefore extremely skewed to the right. In Table 2, distributions were selected so that no one cost element dominates the total. In Table 3, the price and quantity of the first three cost elements were each multiplied by 10 so that the resulting costs comprise a significant portion of the total. In Table 4, only three cost elements were retained, and the price and quantity of the first one was multiplied by 10 so that it dominates the total. All cost elements are assumed to be independent. The data in Table 2 provide the most promising conditions for application of the CLT (the only question is whether 20 cost elements are enough). Because a few extremely right-skewed costs dominate the total in Table 3, we might not expect the CLT to apply here. We might not expect the CLT to apply at all to Table 4, because the first cost element dominates the total and there are only three cost elements.

Figures 6 through 8 show the results of fitting a Beta and a Normal distribution to the mean and variance of total cost shown at the bottom right-hand corners of Tables 2 through 4. The reference point for each graph is the empirical distribution function derived from a simulation of the WBS (a description of this calculation will be given in a later section). Notice that the Normal distribution fits the empirical c.d.f. very well for the first two cases. Even in the third case, the total-cost distribution is not drastically skewed and the Normal distribution provides a respectable fit. This further suggests that the assumption of a very skewed total-cost distribution may not always be valid.

These results beg the question: How can the total-cost distribution be approximately Normal (and hence symmetric) when historical experience shows that total cost is frequently underestimated? Although there are many reasons for cost underestimation, some of the more prominent ones are:

- changes in requirements during the acquisition process,
- occurrence of unforeseen events,
- inadequate allowance for schedule or technical risk,
- overly optimistic component cost estimates from contractors,
- funding instability, and

---

[8] We assume independence for ease of computation and exposition. In general, however, price and quantity will not be independent; if they were, there would be little advantage to modeling them separately. The exact formula for the variance of a product [Mood, Graybill, and Boes (1974)] depends on higher-order moments that few analysts will be able to elicit from expert opinion. In this case, the analyst should employ the propagation-of-error formula to obtain an approximation to the variance of a product of correlated terms.

- political considerations.

Suppose both the estimated and actual distributions of total cost risk are symmetric as depicted in Figure 9. The unknown actual risk distribution properly accounts for the effects of all uncertain events, whereas the estimated risk distribution is limited by lack of knowledge or improper incorporation of some uncertain events into the risk analysis. Consequently, the estimated risk distribution will be centered to the left of and have smaller variance than the actual risk distribution. However, the final program cost will be a random draw from the actual, not estimated, distribution. The result is a much higher probability of exceeding the mean estimated cost than of falling below it. This could lead to the possibly mistaken impression that the distribution of total cost risk is skewed to the right.

The foregoing discussion illustrates that the total-cost distribution is not necessarily skewed to the right just because major program costs tend to be underestimated. The fact that many CERs exhibit log-normal error terms (the log-normal distribution is skewed to the right) indicates that at least one of the assumptions underlying the CLT is often violated in practice. In particular, there may be a high degree of dependency among cost elements, and a few cost elements with the greatest inherent risk may dominate the total program cost. Nevertheless, the cost analyst should exercise caution in assuming a log-normal error term; in fact, this assumption should be examined empirically and tested, if possible.

Table 2. Hypothetical Work Breakdown Structure – Example 1

Price			Quantity			Component Moments					
Low	Mode	High	Low	Mode	High	E(P)	E(Q)	V(P)	V(Q)	E(P*Q)	V(P*Q)
3.48	5.34	12.69	4.83	7.55	16.36	7.17	9.58	3.95	6.05	68.67	697.77
4.18	7.06	13.25	2.14	5.87	17.43	8.16	8.48	3.58	10.59	69.22	1,001.57
4.20	6.38	12.89	5.95	10.80	20.36	7.82	12.37	3.40	8.97	96.76	1,100.16
4.21	6.36	15.00	3.86	6.96	15.82	8.52	8.88	5.44	6.42	75.64	929.47
4.43	5.90	12.95	2.01	5.92	16.87	7.76	8.27	3.46	9.89	64.16	866.52
1.21	2.59	11.07	4.73	8.26	17.67	4.96	10.22	4.75	7.46	50.65	715.01
2.74	4.56	11.71	2.23	4.76	14.08	6.34	7.03	3.75	6.49	44.53	470.13
2.71	4.77	11.99	4.95	7.75	18.24	6.49	10.32	3.95	8.18	66.96	797.90
1.09	2.51	8.80	6.00	8.01	16.17	4.13	10.06	2.81	4.84	41.58	380.45
4.12	6.80	15.33	5.88	10.86	19.16	8.75	11.97	5.71	7.51	104.69	1,434.28
2.76	5.61	13.82	4.40	8.30	17.45	7.40	10.05	5.49	7.48	74.35	1,005.44
1.99	3.60	9.89	2.57	6.41	15.30	5.16	8.09	2.90	7.11	41.74	399.69
4.65	6.60	12.83	3.67	6.28	16.30	8.03	8.75	3.04	7.40	70.25	732.18
3.96	5.41	11.69	2.23	7.17	18.96	7.02	9.46	2.81	12.30	66.37	892.41
2.01	4.78	11.54	3.49	6.38	15.40	6.11	8.42	4.00	6.43	51.45	549.47
2.77	4.59	14.38	3.42	7.18	15.75	7.24	8.78	6.50	6.66	63.59	893.43
2.08	4.16	13.83	5.67	10.50	20.35	6.69	12.17	6.55	9.33	81.50	1,450.20
4.98	7.80	15.21	5.35	9.29	18.97	9.33	11.20	4.66	8.19	104.53	1,335.91
1.93	4.67	13.48	2.80	7.43	18.53	6.69	9.59	6.07	10.89	64.16	1,111.77
3.12	4.61	12.14	5.99	10.24	18.79	6.62	11.67	3.89	7.08	77.30	868.52
									Sum	1,378.10	17,632.27

Table 3. Hypothetical Work Breakdown Structure — Example 2

Price			Quantity			Component Moments					
Low	Mode	High	Low	Mode	High	E(P)	E(Q)	V(P)	V(Q)	E(P•Q)	V(P•Q)
34.80	53.37	126.9	48.28	75.51	163.5	71.69	95.79	395.41	605.17	6,867	6,977,675
41.80	70.57	132.5	21.40	58.66	174.3	81.63	84.79	358.26	1,059	6,922	10,015,731
42.02	63.78	128.8	59.46	108.0	203.6	78.23	123.70	340.37	896.85	9,676	11,001,575
4.21	6.36	15.00	3.86	6.96	15.82	8.52	8.88	5.44	6.42	76	929
4.43	5.90	12.95	2.01	5.92	16.87	7.76	8.27	3.46	9.89	64	867
1.21	2.59	11.07	4.73	8.26	17.67	4.96	10.22	4.75	7.46	51	715
2.74	4.56	11.71	2.23	4.76	14.08	6.34	7.03	3.75	6.49	45	470
2.71	4.77	11.99	4.95	7.75	18.24	6.49	10.32	3.95	8.18	67	798
1.09	2.51	8.80	6.00	8.01	16.17	4.13	10.06	2.81	4.84	42	380
4.12	6.80	15.33	5.88	10.86	19.16	8.75	11.97	5.71	7.51	105	1,434
2.76	5.61	13.82	4.40	8.30	17.45	7.40	10.05	5.49	7.48	74	1,005
1.99	3.60	9.89	2.57	6.41	15.30	5.16	8.09	2.90	7.11	42	400
4.65	6.60	12.83	3.67	6.28	16.30	8.03	8.75	3.04	7.40	70	732
3.96	5.41	11.69	2.23	7.17	18.96	7.02	9.46	2.81	12.30	66	892
2.01	4.78	11.54	3.49	6.38	15.40	6.11	8.42	4.00	6.43	51	549
2.77	4.59	14.38	3.42	7.18	15.75	7.24	8.78	6.50	6.66	64	893
2.08	4.16	13.83	5.67	10.50	20.35	6.69	12.17	6.55	9.33	81	1,450
4.98	7.80	15.21	5.35	9.29	18.97	9.33	11.20	4.66	8.19	105	1,336
1.93	4.67	13.48	2.80	7.43	18.53	6.69	9.59	6.07	10.89	64	1,112
3.12	4.61	12.14	5.99	10.24	18.79	6.62	11.67	3.89	7.08	77	869
									Sum	24,609	28,009,814

Table 4. Hypothetical Work Breakdown Structure — Example 3

Price			Quantity			Component Moments					
Low	Mode	High	Low	Mode	High	E(P)	E(Q)	V(P)	V(Q)	E(P•Q)	V(P•Q)
42.02	63.78	128.9	59.5	108.0	203.6	78.23	123.70	340.37	896.85	9,676	11,001,575
4.21	6.36	15.00	3.86	6.96	15.82	8.52	8.88	5.44	6.42	76	929
1.21	2.59	11.07	4.73	8.26	17.67	4.96	10.22	4.75	7.46	51	715
									Sum	9,803	11,003,219

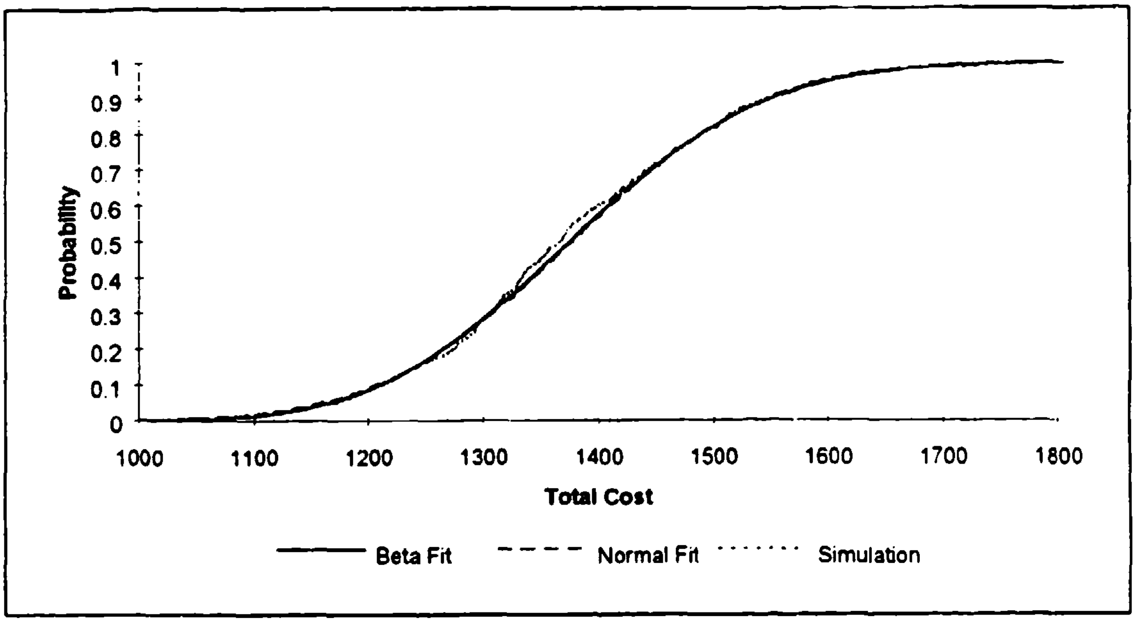

Figure 6. Simulated and Fitted C.D.F.s for WBS in Table 2

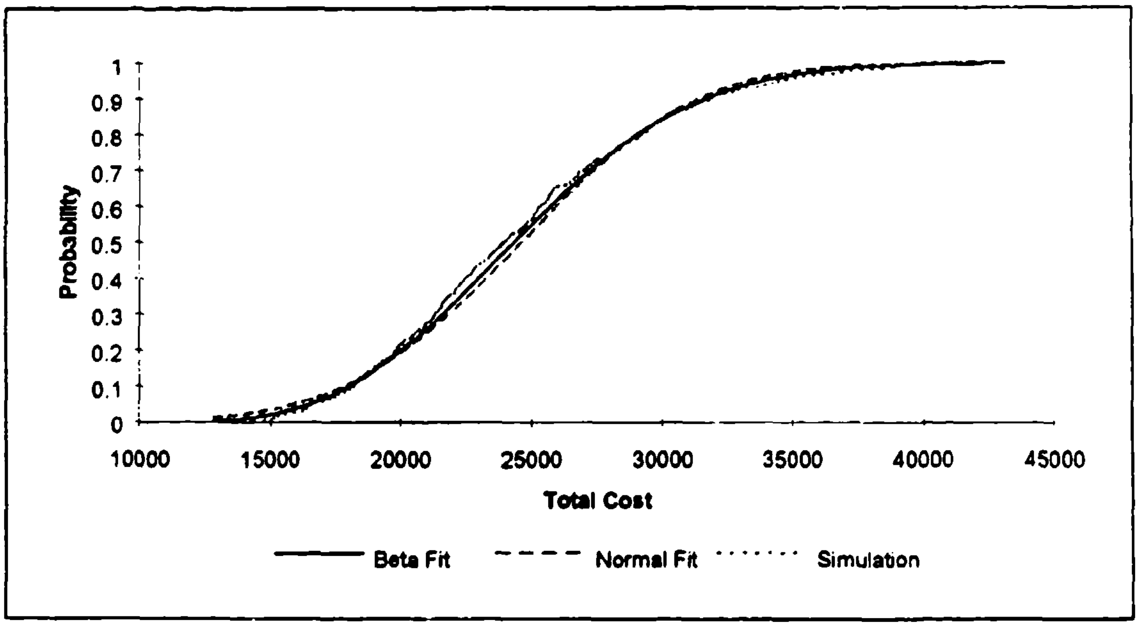

Figure 7. Simulated and Fitted C.D.F.s for WBS in Table 3

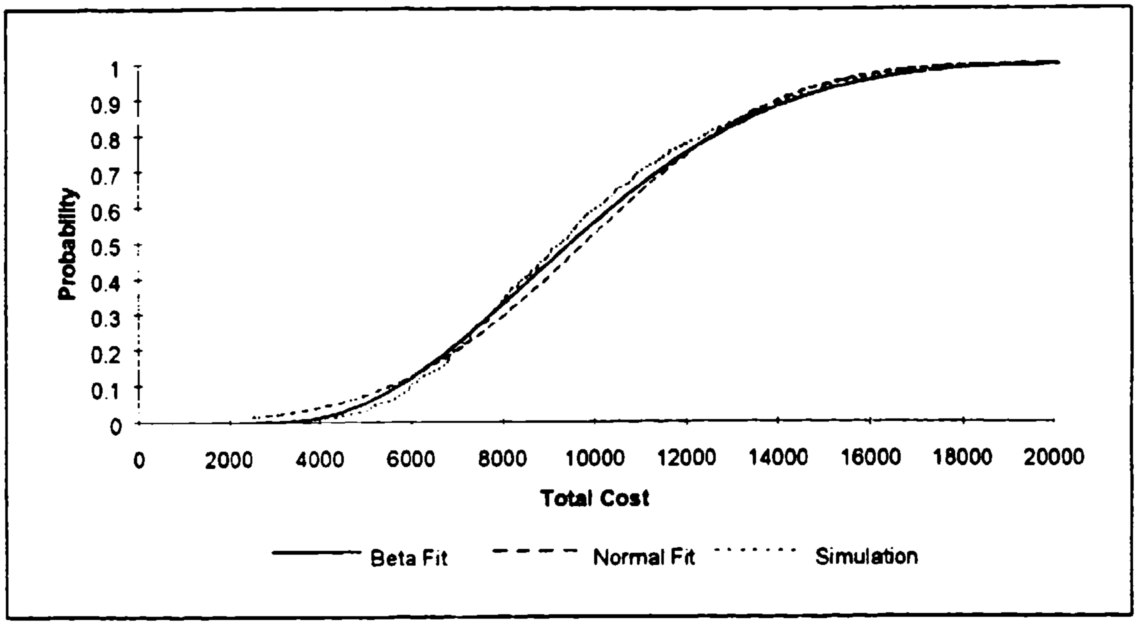

Figure 8. Simulated and Fitted C.D.F.s for WBS in Table 4

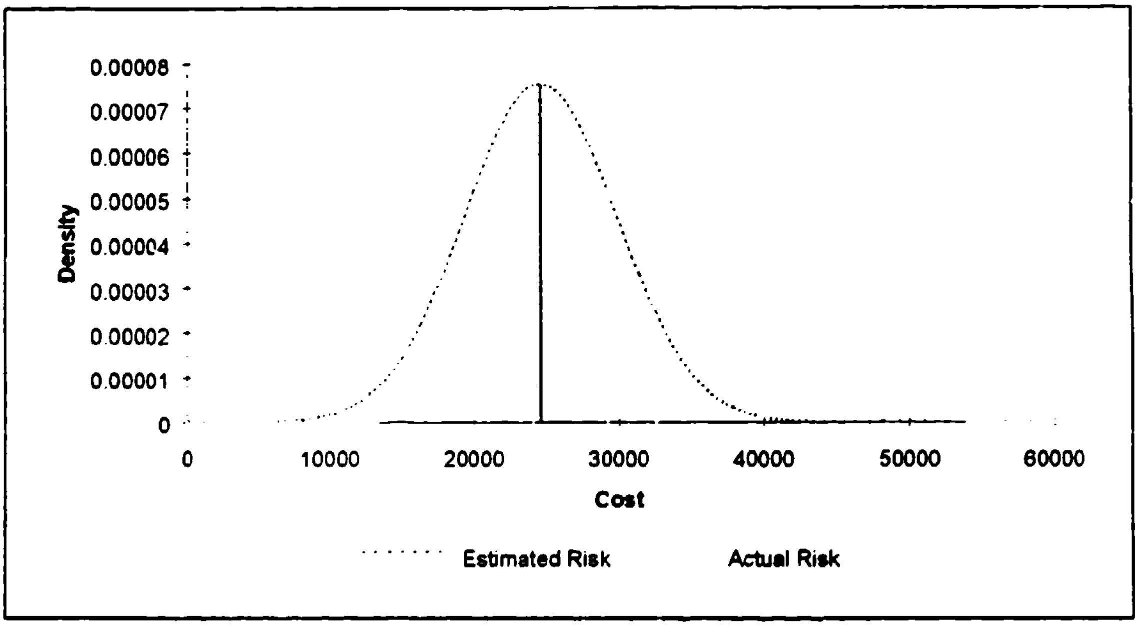

Figure 9. Hypothetical Estimated and Actual Cost Risk Distributions

2. Cost-Element Distributions Determined From CERs

When the cost-element distributions are determined from CERs, the method for obtaining the approximate distribution for total cost is a simple extension of the propagation-of-error technique described in the previous chapter. As a simple example, consider a WBS with two cost elements defined by CERs as follows:

$$Cost_1 = 2x_1^{1.7}x_2^{.6}\varepsilon_1 ,$$

and

$$Cost_2 = 3x_1^{1.2}x_2^{.8}\varepsilon_2 .$$

Then the equation for the total cost is

$$Cost = 2x_1^{1.7}x_2^{.6}\varepsilon_1 + 3x_1^{1.2}x_2^{.8}\varepsilon_2 .$$

First we need to compute all the first- and second-order partial derivatives of the total cost evaluated at the means of the cost drivers and error terms. We then use the propagation-of-error formulas (2) and (3) to compute the approximate mean and variance of the total cost, and fit a Beta distribution to these moments. Note that if the cost elements are correlated, then ε_1 and ε_2 are correlated and must be accounted for in the propagation-of-error formulas. If there are many cost elements in the WBS, this procedure can involve a substantial amount of computation.

B. SIMULATION METHODS

1. Cost-Element Distributions Specified Directly

The method for obtaining the total-cost distribution is straightforward when the cost elements to be simulated are independent. Each cost element realization is generated as a random draw from its specified distribution. After all cost elements have been generated, they are summed to obtain a single realization for the total cost. This process is repeated many times (at least 500 but possibly more depending on the number and variability of the cost elements in the WBS) to obtain a sample from the unknown distribution of total cost. The c.d.f. of total cost is then estimated as

$$F(x) = \frac{1}{n}\sum_{i=1}^{n} I(x - x_{(i)}) , \tag{12}$$

where $x_{(i)}$ is the ith ordered total cost (i.e., the ith value of total cost when sorted in ascending order) and

$$I(y) = \begin{cases} 1 \text{ if } y \geq 0 \\ 0 \text{ if } y < 0 . \end{cases}$$

A problem with using simulations to estimate total cost risk is the necessity to rerun the entire simulation when even a single input parameter has changed. It can therefore be extremely time-consuming to perform sensitivity analyses in conjunction with simulation. One possible shortcut is to fit a response surface by regressing the simulation output variable against a set of key input variables [Iman and Helton (1988)]. The value of the output variable under a new set of inputs can then be approximated by a point on the response surface. This approach is not a panacea, however, because the response surface may not provide a good fit to the computer model. Moreover, a smooth response surface may mask discontinuities in the true underlying relationship between simulation inputs and outputs. It may be better to simply run the simulation over when the inputs change.

The biggest drawback to performing simulations is the lack of an acceptable method for generating correlated non-normal random variables with bounded domains (such as Beta and Triangular distributions). Assuming the cost-element distributions are functionally related (i.e., complete positive dependence among all cost elements) will give an upper bound on the variance of the total-cost distribution. Similarly, assuming complete negative dependence among all cost elements will give a lower bound on the variance of total cost. However, the range of distributions between these two bounds is often so large as to make this exercise meaningless. Until an acceptable method for simulating correlated cost elements is available, the analytical (i.e., propagation-of-error) method or a method based on CERs (see the next section) is preferable.

2. Cost-Element Distributions Determined From CERs

When estimates of cost and risk are needed at lower levels of aggregation as well as for the total, one approach is to model cost-element subtotals with CERs and then aggregate them. Another situation in which cost elements in a WBS might be modeled with a CER is when the analyst believes there is dependence among the cost elements. Dependent cost elements would then be modeled to have at least one explanatory variable in common. The last chapter described the method for simulating a cost determined from a CER. The only difference when simulating cost elements in a WBS is that many CERs are involved. Once each cost element in the WBS has been generated from its CER, the procedure for determining the total-cost distribution is the same as when the cost-element distributions are specified directly.

IV. STOCHASTIC NETWORK MODELS

Stochastic network models take work breakdown structures to a higher level of sophistication by allowing a detailed description of the actual processes involved in determining the cost, schedule, and performance of a particular program or system. A network consists of arcs and nodes, the arcs representing program activities and the nodes representing decision points. Not only does the network represent the program's activities and decision points, but the phasing of them as well. A network is also characterized by its nodal logic, which determines what actions are taken on input and output to each node. Because the actual events in the program life cycle—with attendant cost, schedule, technical risk, and other consequences—are modeled, the analyst can trace influential sources of risk and take steps to mitigate them. This is the primary advantage of the network model over CERs and work breakdown structures.

The network itself is merely a detailed "wiring diagram" of the process activities, interactions, and decision mechanisms. A stochastic network allows randomness to enter the description of activities and decision points. For example, the cost for a particular activity may be specified as a random variable with a Beta distribution. Similarly, the schedule for that activity can be modeled as a random process as well. At decision points, probability distributions can specify the possible paths to take in the network. For example, if the integration of the avionics, airframe, and propulsion system for an aircraft under development is successful, then the system will proceed as planned. If, however, the integration is unsuccessful, some redesign is likely in order and the network will branch back to a redesign phase. This will result, of course, in a schedule delay and increased cost.

The primary drawback to a stochastic network model is the work involved in putting it together. It usually takes the collaborative efforts of many managers, engineers, contractors, and other analysts to construct one because there is rarely one person who understands every detail of the activities in the network. Rather, an individual may be able to construct a sub-network for the activities with which he/she is familiar, and the program manager or someone familiar with the overall picture of how the activities fit together will link the sub-networks together. This can take many man-months of effort to construct.

There is also the danger in disaggregating a network too far. Some analysts believe that the more disaggregated and detailed a breakdown, the better. We disagree with this assertion in general. The network should be carried to a level of detail for which reasonably accurate assessments of cost, schedule, and technical risk can be made. If only rough estimates of risk are required, a relatively aggregate representation of the network may be sufficient. (Although if only rough preliminary estimates are needed, it hardly seems worth the effort to construct a network in the first place. A CER or WBS may be all that is needed.) On the other hand, if a network is disaggregated to a level of detail for which accurate information or knowledge is unavailable, then the analyst risks introducing more noise into the system. Exactly how much detail is appropriate to accurately model risk is an open question, the answer to which likely depends on the individual circumstances being modeled.

We illustrate the use of stochastic networks with a very simple example. The example, taken from Mann (1979) and reproduced in Figure 10, displays a very aggregate network representation of the development of a hypothetical fighter aircraft.

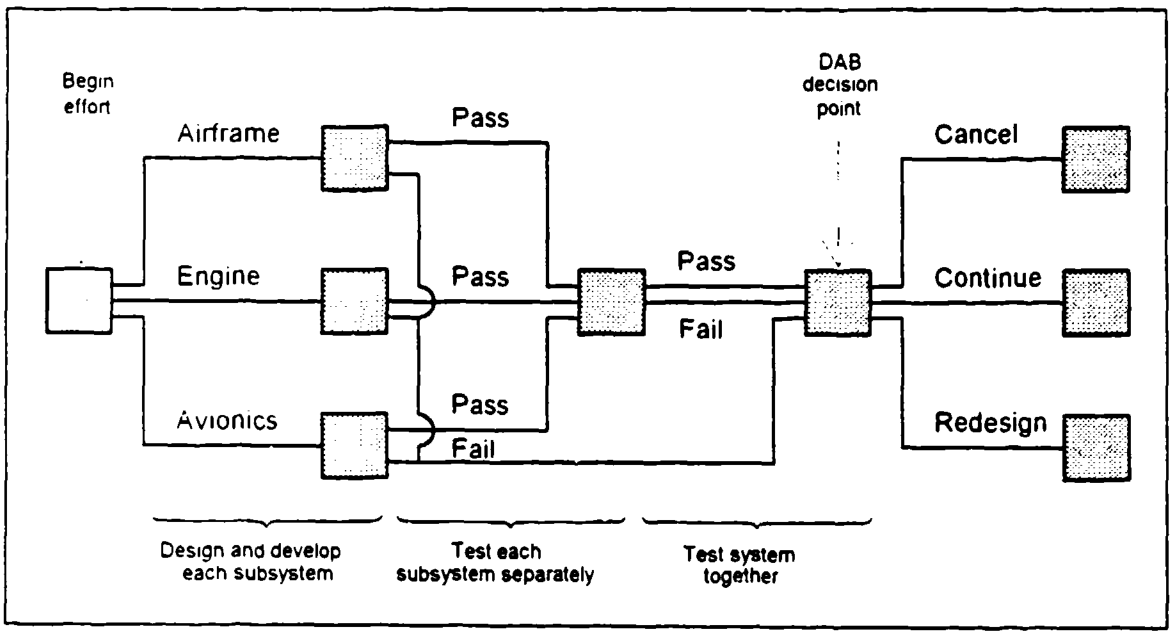

Figure 10. Example of Network for Development of Hypothetical Fighter Aircraft

The process begins with the parallel development of the airframe, engine, and avionics. If all three of these are successfully developed to specification, the next stage is to integrate them into a complete system. If any single development effort fails, a decision must be made about how to proceed. The decision could be either cancellation of the project, continuation of the project with reduced requirements, or a redesign of the unsuccessful components. The same types of decisions must be made if the subsystems cannot be successfully integrated into a functional unit.

In actuality, of course, the development process is much more complex. There will be many more branches and decision points in the system and more complex nodal logic will be required. Even the simple network depicted in Figure 10 can be made much more complicated by allowing more complex logic to apply. For example, if all three subsystems must be completed prior to integration, the time elapsed until the beginning of the integration phase is the maximum of the three subsystem development times. In this case, the probability of reaching the integration phase by time t is:

$$F(t) = P(T < t) = F_1(t)F_2(t)F_3(t) \,,$$

where $F_1(t)$, $F_2(t)$, and $F_3(t)$ are the probabilities of completion before time t for the three subsystems. If, on the other hand, only two subsystems need be completed prior to integration, the probability of reaching the integration phase by time t is:

$$F(t) = F_1(t)F_2(t)(1 - F_3(t)) + F_1(t)F_3(t)(1 - F_2(t)) + F_2(t)F_3(t)(1 - F_1(t)) + F_1(t)F_2(t)F_3(t).$$

As demonstrated by this simple illustration, the mathematics involved in the computation of network probabilities can rapidly become cumbersome with the introduction of only a few twists in the logic. In a real application, the mathematics are almost always intractable, and simulation is required to compute the total schedule and cost risk of the system.

We will give a very simple illustration of how to simulate a network, referring to the hypothetical network depicted in Figure 10. There are usually many output quantities of interest in a network model, for example, the probability of successful program completion, the total program cost, the probability of exceeding the schedule by a certain percentage, the expected times and costs to reach program milestones, and so on. In this example, we compute the probability distribution of total program cost and the probability that the system will need to be redesigned. To do this, we made assumptions regarding the stochastic nature of the network activities and events, as shown in Table 5.

Each simulation of the network begins with the parallel development of the airframe, engine, and avionics. However, the costs of these subsystems will vary from iteration to iteration. The network will also branch differently for each iteration based on the probabilities shown in Table 5. Integration costs will be incurred only if all three subsystem tests are successful. If we tabulate the paths taken (with associated

Table 5. Probabilities Associated With Network Activities and Events

Activity/Event	Distribution[a]	Parameters		
Airframe Cost	Triangular	$l = 100$,	$m = 150$,	$h = 250$
Engine Cost	Triangular	$l = 75$,	$m = 125$,	$h = 175$
Avionics Cost	Triangular	$l = 200$,	$m = 300$,	$h = 500$
Integration Cost	Triangular	$l = 150$,	$m = 200$,	$h = 275$
Successful Airframe Test	Bernoulli		$p = .90$	
Successful Engine Test	Bernoulli		$p = .85$	
Successful Avionics Test	Bernoulli		$p = .70$	
Successful Integration Test	Bernoulli		$p = .95$	
DAB Decision if Subsystem Fails[b]	Trinomial	$p_1 = .25$,	$p_2 = 0$,	$p_3 = .75$
DAB Decision if Integration Succeeds	Trinomial	$p_1 = .02$,	$p_2 = .95$,	$p_3 = .03$
DAB Decision if Integration Fails	Trinomial	$p_1 = .10$,	$p_2 = .05$,	$p_3 = .85$

[a] The Triangular distribution is characterized by a lowest (l), modal (m), and highest (h) cost. The Bernoulli distribution is used to model a binary event (for example, success/failure) with probability of success p. The Trinomial distribution is used to model an event with three possible outcomes. In our notation, p_1 is the probability of cancellation, p_2 is the probability of continuation, and p_3 is the probability of redesign.

[b] Alternatively, the decision to proceed may be modeled deterministically, for example, the program may be canceled if the cost at the decision point exceeds a specified value.

costs) through the network for each iteration, at the end of the simulation we will be able to compute the percentage of times a particular node was reached and the distribution of total cost.

Because the network in this example is so small, we tabulated the paths taken through the network for each iteration using Microsoft *Excel* and an add-in simulation package called *@RISK*. In general, a true network simulation will require special-purpose software to run (some software options are discussed in the next chapter). After simulating 1,000 passes through the network, we computed the distribution of total cost and the probabilities of cancellation and redesign. The distribution of total cost was calculated using equation (12) and is displayed in Figure 11. The probabilities of cancellation ($p = .126$) and redesign ($p = .375$) were calculated by simply tabulating the percentage of times those decisions were reached during the network simulation.

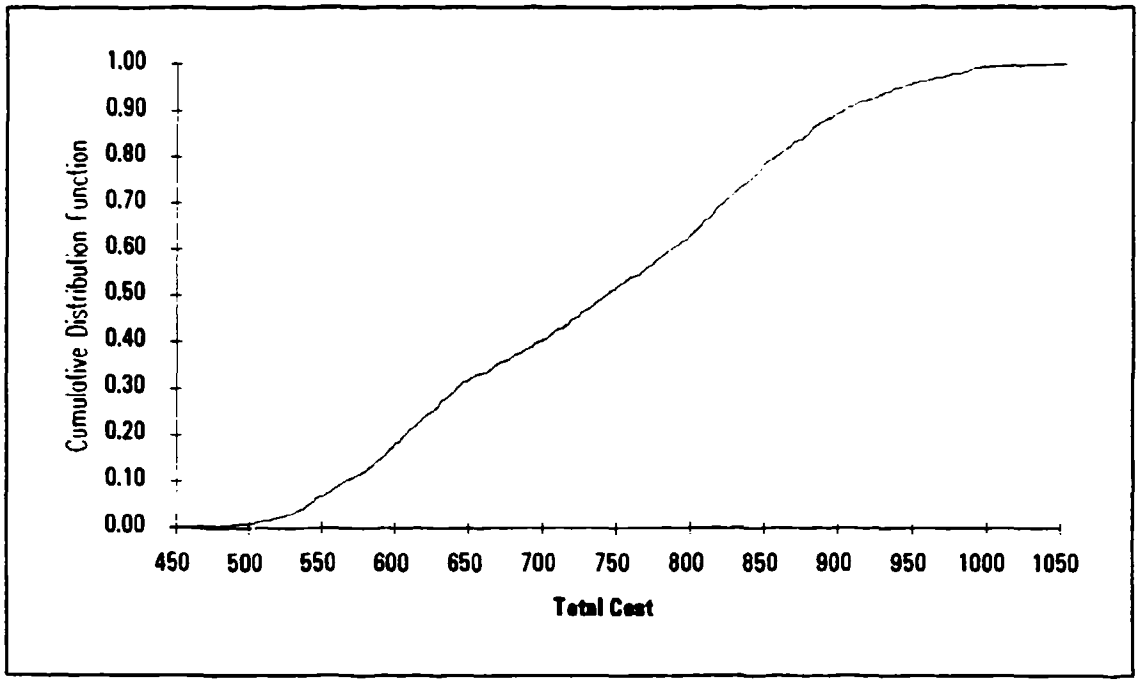

Figure 11. Estimated Distribution of Total Cost From Network Simulation

This example has illustrated, under greatly simplified conditions, how stochastic networks can be used to assess total cost and schedule risk. In deciding whether to use a stochastic network to model an actual defense program, the analyst will have to decide whether it is worth the time and effort to construct the network, with its associated nodal logic and probability distributions, for this is where most of the difficulty is involved. Furthermore, we are unaware of any evidence to indicate that stochastic network models yield more accurate assessments of risk than simpler methods, such as CERs and work breakdown structures. It would therefore seem that stochastic networks are most useful in situations where it is necessary to pinpoint and control specific sources of risk rather than simply measure the risk.

APPENDIX

SOFTWARE FOR RUNNING SIMULATIONS

APPENDIX

SOFTWARE FOR RUNNING SIMULATIONS

In cost risk analysis, the objective is often to determine the probability distribution of an algebraic combination of random variables. For example, total system cost may be computed as the sum of the costs of all subsystems. If more detail is available, subsystem cost may be computed as the sum of the costs of all components, and so on.

Particular cost elements are often expressed as the product of a price per unit and the number of units. For example, the cost of a particular grade of production labor may be expressed as the product of the hourly wage and the number of manhours. Similarly, the cost of a particular type of production material may be expressed as the product of the price per unit (e.g., per pound or per square foot) and the number of units. In all of these cases, the probability distributions are assumed known at the lowest level of aggregation. The objective is to combine the individual probability distributions to obtain the probability distribution of total system cost.

It is often extremely difficult to obtain a mathematical expression for the probability distribution of sums or products of random variables, particularly when they are correlated [Springer (1979)]. Consequently, recourse is often made to simulation. During a single iteration, the underlying random variables are each drawn from their respective distributions. The algebraic combination of these random variables, representing total system cost, is then computed (price and quantity are multiplied, component costs are added, etc.). This process is repeated for many hundreds or even thousands of iterations, and a record is maintained of total system cost for each iteration. After the desired number of iterations have been performed, the probability distribution of total system cost is analyzed. In particular, the values of this variable are sorted, and key percentiles are computed (e.g., 50th, 90th, and 95th percentiles). Summary statistics such as the mean, mode, and variance are also computed.

The software to perform cost risk analysis differs according to the method used: cost-estimating relationships (CERs), work breakdown structure (WBS), and network modeling. When using CERs or a WBS, the cost risk analysis is often performed inside a spreadsheet. The most useful software packages in these cases are add-ins to popular commercial spreadsheets. On the other hand, several stand-alone software packages are available for network modeling.

A. COST RISK ANALYSES BASED ON CERS AND WBS

By estimating a CER, the analyst obtains estimates of the regression parameters, their standard errors, and the standard deviation of the error term. It is implicitly assumed in the estimation procedure that the error terms are normally distributed, possibly after a transformation. For example, equation (4) might be estimated by regressing the logarithm of cost against the logarithms of the cost drivers, x_1 and x_2. If the regression is performed on the logarithms of the variables, it is assumed that the error in predicting the *logarithm* of cost is normally distributed. If so, then the error in predicting the *level* of cost is log-normally distributed.

In addition, the analyst is presumed to have estimates of the probability distributions of the cost drivers. Simulation methods thus differ from analytical methods (such as propagation of error), in that the latter require only the means, variances, and correlations of the cost drivers, not their fully-specified distributions. If the analysis is based on a WBS rather than a CER, then the analyst directly specifies the probability distribution of each cost element. The probability distributions may derive from expert opinion or, if there are sufficient data, from statistical estimation.

1. Spreadsheet Add-in Software

The two most widely-used commercial spreadsheets are *LOTUS 1-2-3* and Microsoft *Excel*. A popular add-in to these spreadsheets is the *@RISK* package.[9] We have evaluated *@RISK* version 1.1 for Microsoft *Excel*, dated 6 February 1992. A competing add-in package is *Crystal Ball for Windows*.[10] We have evaluated version 2.0 dated 1 December 1991.

For the most part, the two add-in packages operate similarly. Using either package, any cell in a spreadsheet may be defined as a probability distribution. For example, one cell might represent the hourly wage of a particular grade of production labor, and another cell might represent the number of manhours. The hourly wage might be defined as a normal distribution with mean 20 and variance 16. The number of

[9] Palisade Corporation, 31 Decker Road, Newfield, New York 14867, telephone 1-800-432-RISK.

[10] Decisioneering, Inc., 1727 Conestoga Street, Boulder, Colorado 80301, telephone 1-800-289-2550.

manhours might be defined as a triangular distribution with a low value of 1,000, a mode of 2,000, and a high value of 5,000.

In fact, *@RISK* allows the following probability distributions: beta, binomial, chi-squared, discrete uniform, Erlang, exponential, gamma, geometric, hypergeometric, logistic, log-normal, negative binomial, normal, Pareto, Poisson, triangular, uniform, and Weibull. By contrast, *Crystal Ball* allows the following probability distributions: beta, binomial, exponential, geometric, hypergeometric, log-normal, normal, Poisson, triangular, uniform, and Weibull.

All *Excel* arithmetic operations may be performed on cells defined as probability distributions. In our example, a new cell representing labor cost may be defined as the product of the hourly wage and the number of manhours. The probability distributions specified for the latter two variables imply a probability distribution for labor cost. With each simulation of the spreadsheet, new values are drawn for the hourly wage and the number of manhours, and a new value (i.e., the product) is computed for labor cost.

Until the simulation option is invoked, the cells display only a single value for each variable. Depending on the situation, this value may be the mean of the specified probability distribution, or may be the realization of the most recent simulation. The simulation option is invoked by clicking the mouse on a simple command from the menu bar. At that point, values are drawn from each probability distribution in the spreadsheet. All other cells are then recomputed, including those linked by arithmetic operations to the simulated cells. Increments are made to running counters of specified "results" cells. These steps comprise a single iteration of the simulation.

The entire procedure is repeated for a number of iterations specified by the user. After all iterations have been computed, the user may construct and view a variety of reports. These include histograms, cumulative distribution curves, summary graphs, and statistical reports. In addition, the user may request the probability that result cells fall above or below specified target values.

Neither package allows the user to specify Pearson (i.e., product-moment) correlations between simulated variables. However, both packages allow the user to specify the less-familiar Spearman (i.e., *rank*) correlations. Using either package, the user may directly specify the rank correlation between any pair of cells, or the rank correlation matrix corresponding to any range of cells. Alternatively, if the user provides two columns of data, *Crystal Ball* will estimate the rank correlation and apply the result to a specified pair of cells.

2. Microsoft *Excel* Version 4.0

The latest release of Microsoft *Excel* has an extensive simulation capability, reducing the need to rely on add-in software. The "Analysis Tools" menu includes an option for random number generation. Any cell in a spreadsheet may be defined as one of the following probability distributions: Bernoulli, binomial, normal, Poisson, or uniform.

Another approach is possible if this list of distributions is inadequate. A random number from a given probability distribution may be generated by transforming a uniform random number. Specifically, the inverse cumulative distribution function of the desired distribution is applied to the unifrm random number. Under the "Statistical Functions" menu, *Excel* provides the following inverse cumulative distribution functions: beta, chi-squared, gamma, log-normal, nonnal, and t-distribution.

An apparent omission is that *Excel* does not provide the inverse of the triangular distribution function. However, that inverse is easily obtained in closed-form. Consider the triangular distribution with low value l, modal value m, and high value h. The cumulative distribution function is piecewise quadratic:

$$F(x) = \begin{cases} (x-l)^2/[(h-l)(m-l)], & \textit{for } l < x < m \\ 1-\{(h-x)^2/[(h-l)(h-m)]\}, & \textit{for } m < x < h \end{cases}.$$

Inversion of this function involves solving a quadratic equation, admitting the possibility of two distinct roots. However, elementary analysis reveals that only one of the two roots falls between the low and high values. After obtaining the uniform random number, u, the triangular random number is obtained by the following transformation:

$$\text{If } 0 < u < (m-l)/(h-l), \quad \text{set } x = l + \sqrt{(h-l)(m-l)u}$$

$$\text{If } (m-l)/(h-l) < u < 1, \quad \text{set } x = h - \sqrt{(h-l)(h-m)(1-u)}.$$

B. COST RISK ANALYSES BASED ON NETWORK MODELS

Many stand-alone software packages are available for performing network modeling. One such package, Venture Evaluation and Review Technique (VERT), is particularly well-documented [Lee et al. (1982), Moeller and Digman (1981)]. VERT associates three random variables with each arc in the network: cost incurred, time elapsed, and performance generated. Note that performance refers to the manufacturing process being modeled, not the end-item that results from that process (e.g., amount of rework performed, as opposed to the speed or maneuverability of the final aircraft.) The cost, time, and performance variables may be modeled using the following probability distributions: beta, binomial, chi-squared, Erlang, gamma, geometric, hypergeometric, log-normal, normal, Poisson, triangular, uniform, and Weibull.

VERT allows three main logical structures for its nodes. These are described in Lee et al. (1982):

- AND input logic requires that *all* the input arcs to the node be successfully completed before the combined input network flow is transferred to the output logic for the appropriate distribution among the output arcs. The time computed for the models bearing AND input logic is the maximum cumulative time of all the input arcs. Cost and performance are computed as the sum of all the cumulative cost and composite performance values of all input arcs.

- PARTIAL AND input logic is similar to AND input logic, except that it requires a minimum of one input arc to be successfully completed before allowing flow to continue through this node. However, this logic will wait for all the input arcs to come in or to be eliminated from the network before processing. The calculations of node time, cost, and performance values for this input logic are the same as those used for the AND logic.

- OR input logic requires a minimum of just one input arc to be successfully completed before allowing the flow to continue through this node. This logic, unlike PARTIAL AND logic, will not wait for all the input arcs to come in or to be eliminated from the network before the flow is processed. Therefore, as soon as an input arc is successfully completed, the flow will be sent on to the output logic for processing. The time and performance assigned to the node are the cumulative time and performance values carried by the first successful input arc to be processed. Cost is computed as the sum of all the cumulative costs of all the active input arcs.

Although these definitions require contemplation, a simple example illustrates our understanding of the differences among them. Consider again an aircraft composed of an airframe, engine, and avionics. The AND logic applies if system integration cannot proceed until *all three* major subsystems have been completed. System integration begins at the largest completion time of the three subsystems (i.e., the latest arrival). Cost and performance are aggregated over all three subsystems.

To illustrate PARTIAL AND logic, suppose that maintenance manuals are written for each subsystem after design and manufacture of that subsystem have been completed. It is possible to begin writing after the first subsystem has been completed (i.e., the earliest arrival). Writing of the second and third manuals will be delayed until those subsystems have been completed. Although the three manuals are staggered in time, cost and performance are again aggregated over all three subsystems.

Finally, to illustrate OR logic, suppose that two alternative engine designs are being considered, and that the first to be successfully completed will be adopted. Engine design will stop at the earliest arrival time. Performance will be measured for this earliest arrival, because only this design is adopted. Cost, however, is aggregated over both design processes, *through the date of the earliest arrival.* There is no need to consider costs that would have been incurred to complete the alternative design, because that design process is abandoned.

REFERENCES

REFERENCES

Abeyta, Richard, Paul Grover, and Monte Norton. "Production Risk Assessing Methodology (PRAM)." U.S. Army Procurement Research Office. May 1982.

Abramson, Robert L., and Philip H. Young. "Using Risk-Impact Drivers to Form WBS-Element Cost Means and Variances." The Aerospace Corporation, September 1991.

Abramson, Robert L., and Philip H. Young. "FRISKEM — Formal Risk Evaluation Methodology." The Aerospace Corporation, October 1990.

Abramson, Robert L., and Stephen A. Book. "A Quantification Structure for Assessing Risk-Impact Drivers." The Aerospace Corporation, 1990.

Army Armament Command. "Cost/Schedule Uncertainty Analysis of the XM1/Alternative Armament Programs." April 1976.

Batson, Robert G. "Cost Risk Analysis Methodology: A State-of-the-Art Review." *Estimator*, Spring 1989.

Black, R. L. and J. J. Wilder. "Probabilistic Cost Approximation When the Inputs are Dependent." Grumman Aerospace Corporation, PDR-OP-B82-01, April 1982.

Borison, Adam B., and Thomas R. Rice. "Probabilistic Cost Estimating for the Great Plains Coal Gasification Plant." *Interfaces*, Vol. 11, No. 2, April 1981.

Budescu, David V., and Thomas S. Wallsten. "Encoding Subjective Probabilities: A Psychological and Psychometric Review." *Management Science*, Vol. 29, No. 2, February 1983.

Chapel, Stephen W., Edward W. Merrow, and Christopher Worthing. "A Review of Cost Estimation in New Technologies: Implications for Energy Process Plants." The RAND Corporation, R-2481-DOE, July 1979.

Clark, Charles E. "The *PERT* Model for the Distribution of An Activity Time." *Operations Research*, Vol. 10, 1962.

Cooke, R. M. *Experts in Uncertainty, Opinion and Subjective Probability in Science*. New York: Oxford University Press, 1991.

Cooper, Richard, and Gary Nelson. "Analytic Methods for Adjusting Subjective Rating Schemes." The RAND Corporation, R-1685-ARPA, 1976.

Cox, D. C., and P. Baybutt. "Methods for Uncertainty Analysis: A Comparative Survey." *Risk Analysis*, Vol. 1, No. 4, 1981.

Defense Systems Management College. "Risk Assessment Techniques: A Handbook for Program Management Personnel." July 1983.

Defense Systems Management College. "Risk Management." 1990.

Department of the Air Force, HQ AFSC/ACC. "AFSC Risk Model." August 15, 1985.

Devaney, R., and P. Popuvic. "An Evaluation of Component Dependence in Cost Risk Analysis." Air Force Institute of Technology, AFIT/GSM/LSQ/85S-27, 1985.

Dienemann, P. F. "Estimating Cost Uncertainty Using Monte Carlo Techniques." The RAND Corporation, RM-4854-PR, January 1966.

Dyer, John L. "System Cost Probability Distributions for ACES Architectures." Sparta, Inc., May 30, 1989.

Edgar, Lt. Col. John D. "Controlling Murphy: How to Budget for Program Risk." *Software Risk Management*. Washington, D.C.: IEEE Computer Society Press, 1989.

Feller, William. *An Introduction to Probability Theory and Its Applications*. Vol. II, Second Edition. New York: John Wiley & Sons, Inc., 1971, pp. 262-63.

Fischoff, B., P. Slovic, and S. Lichtenstein. "Improving Intuitive Judgments by Subjective Sensitivity Analysis." Decisions and Designs, Inc., Report PTR-1060-78-7, July 1978.

Free, L. J. "Cost/Schedule Growth Prediction Methodology." Naval Underwater Systems Center, Report 8435, 1989.

Garvey, Paul R. "A General Analytic Approach to System Cost Uncertainty Analysis." In *Cost Analysis and Estimating: Tools and Techniques*, Willis R. Greer, Jr. and Daniel A. Nussbaum (eds.). New York: Springer-Verlag, 1990.

Garvey, Paul R. "Measuring Uncertainty in a Software Development Effort Estimate." *Journal of Parametrics*, Vol. VI, No. 1, March 1986.

Garvey, Paul R., and Frederic D. Powell. "Three Methods for Quantifying Software Development Effort Uncertainty." *Software Risk Management*. Washington, D.C.: IEEE Computer Society Press, 1989.

Goicoechea, A., M. Kraus, and L. Antle. "An Approach to Risk and Uncertainty in Benefit-Cost Analysis of Water Resource Projects." U.S. Army Corps of Engineers, Institute for Water Resources, VA, 1982.

Howard, Truman W., III. "Methodology for Developing Total Risk Assessing Cost Estimate (TRACE)." U.S. Army Missile Research and Development Command, November 1978.

Iman, Ronald L., and Jon C. Helton. "An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models." *Risk Analysis*, Vol. 8, No. 1, 1988.

Johnson, Norman L. and Samuel Kotz. *Distributions in Statistics: Continuous Univariate Distributions - 2*. New York: John Wiley & Sons, Inc., 1970, pp. 37-56.

Jordan, Harley R., and Michael R. Klein. "An Application of Subjective Probabilities to the Problem of Uncertainty in Cost Analysis." Office of the Chief of Naval Operations, Resource Analysis Group, November 1975.

Kahneman, D., P. Slovik, and A. Tversky. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press, 1982.

Keefer, D. L., and S. E. Bodily. "Three-Point Approximations for Continuous Random Variables." *Management Science*, Vol. 29, No. 5, May 1983, pp. 595-609.

Kerns, W. "Risk and Uncertainty: State-of-the-Art in Application." In *1982 Federal Acquisition Research Symposium. Integrating Theory and Experience: The Acquisition Research Connection*, 1982.

Lee, Sang M., Gerald L. Moeller, and Lester A. Digman. *Network Analysis for Management Decisions: A Stochastic Approach*. Boston, MA: Kluwer-Nijhoff Publishing, 1982.

Li, Shing Ted, and Joseph L. Hammond. "Generation of Pseudo-Random Numbers With Specified Univariate Distributions and Correlation Coefficients." *IEEE Transactions on Systems, Man, and Cybernetics*, September 1975.

Lieber, R. "New Approaches for Quantifying Risk and Determining Sharing Arrangements." In *1982 Federal Acquisition Research Symposium. Integrating Theory and Experience: The Acquisition Research Connection*, 1982.

Long, John Amos. "Life Cycle Costing in a Dynamic Environment." Ph.D. Thesis, Ohio State University, 1983.

Lurie, P. M., and M. S. Goldberg. "Simulating Correlated Distributions With Bounded Domains." Institute for Defense Analyses, Paper P-2732, September 1992.

Mallik, A. "Uncertainty Assessment in Life Cycle Cost Analysis." U.S. Army Corps of Engineers, Construction Engineering Research Laboratory, P-85/12, 1985.

Mann, Major Greg A. "VERT: A Risk Analysis Tool for Program Management." *Defense Management Journal*, May-June 1979.

Martin, Col. M. D., Dr. Alan J. Rowe, and Dr. Herold A. Sherman, eds. *Proceedings: Management of Risk and Uncertainty in the Acquisition of Major Programs*. University of Southern California, May 1981.

McNabb, Dale E. "Triangular Frequency Distributions: An Old Concept Updated for the Nineties." In *1989 Federal Acquisition Research Symposium*, 1989.

McNichols, Gerald R. "On the Treatment of Uncertainty in Parametric Costing." Ph.D. Thesis, The George Washington University, February 1976.

McNichols, Gerald R. "Cost Risk Procedures for Weapon System Risk Analysis." In *Proceedings of the Annual Reliability and Maintainability Symposium*, 1981.

Meridan Corporation. "Analysis of Quantitative Techniques for DARPA Program Risk Management." Interim Report, May 19, 1983.

Merrow, Edward W. "Cost Growth in New Process Facilities." The RAND Corporation, April 1983.

Moeller, G... L. and Lester A. Digman. "Operational Planning with VERT." *Operations Research*, Vol. 29, No. 4, July-August 1981.

Mood, Alexander, Franklin Graybill, and Duane Boes. *Introduction to the Theory of Statistics*. Third Edition. New York: McGraw Hill Book Company, 1974, p. 180.

Morgan, M. Granger, and M. Henrion. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge: Cambridge University Press, 1990.

Mosteller, F. "Assessing Unknown Numbers: Order of Magnitude Estimation." In *Statistics and Public Policy*, W. Fairley and F. Mosteller (eds.). Reading: Addison-Wesley, 1977.

Murphy, E. "Statistical Methods of Measuring the Uncertainty of Cost Estimates." U.S. Army Missile Command, 1970.

Myer, M. A., and J. M. Booker. *Eliciting and Analyzing Expert Judgment: A Practical Guide*. London: Academic Press, 1990.

Myers, R. H. *Response Surface Methodology*. Ann Arbor: Edwards Brothers, 1976.

O'Hara, Thomas R. "An Investigation of the Air Force Risk Model." Thesis, Air Force Institute of Technology, September 1991.

Poulton, E. C. *Bias in Quantifying Judgments*. Hillsdale: Lawrence Erlbaum Associates, Publishers, 1989.

Rowe, A. J., I. A. Somers. "Methods to Predict Cost Overruns in the Acquisition Process." In *1982 Federal Acquisition Research Symposium. Integrating Theory and Experience. The Acquisition Research Connection*, 1982.

Schaefer, D. F., F. J. Husic, and M. F. Gutowski. "A Monte Carlo Simulation Approach to Cost Uncertainty Analysis." Research Analysis Corporation, RAC-TP-349, 1969.

Seaver, David A., Ward Edwards, and Detlof von Winterfeldt. "Eliciting Subjective Probability Distributions on Continuous Variables." *Organizational Behavior and Human Performance*, Vol. 21, 1978.

Seiler, F. A. "Error Propagation for Large Errors." *Risk Analysis*, Vol. 7, No. 4, 1987.

Smith, Patrick L. "Quantifying Space System Cost Growth Risk at the Preliminary Design Stage." The Aerospace Corporation, Report ATM 89(1824)-1, September 6, 1989.

Springer, M. D. *The Algebra of Random Variables*. New York: John Wiley & Sons, Inc., 1979.

Sutherland, W. "Fundamentals of Cost Uncertainty Analysis." Research Analysis Corporation, RAC-CR-4, March 1971.

Sutherland, W. "Adding Cost Estimates That Are Not Symmetric About the Most Likely Value." Research Analysis Corporation, RAC-CR-23, 1971.

The Analytic Sciences Corporation. "AFSC Cost Estimating Handbook." Book II, Vol. 1, Chap. 13, 1985.

Thomas, Thomas N. "VERT — A Risk Analysis Technique for Program Managers." Defense Systems Management College. Program Management Course Individual Study Program, Study Project Report PMC 77-1, May 1977.

Timson, F. S., and D. P. Tihansky "Confidence in Estimated Airframe Costs: Uncertainty Assessment in Aggregate Predictions" The RAND Corporation, R-903-PR, October 1972.

Tukey, John W. "The Propagation of Errors, Fluctuations and Tolerances: Basic Generalized Formulas." Princeton University, Technical Report No. 10, December 1957.

Von Winterfeldt, D., and W. Edwards *Decision Analysis and Behavioral Research.* Cambridge: Cambridge University Press. 1986.

W. J. Schafer Associates, Inc. "Life Cycle Cost Inputs, Assumptions and Methods, Cost Appendix," Vol. 1, February 10, 1988.

Whitfield, R. G., and D. E. Newsom. "A Survey of Methods for Uncertainty Analysis." Argonne National Laboratory, April 1984.

Wilder, J. J., and R. L. Black. "Determining Cost Uncertainty in Bottoms-Up Estimating." In *1982 Federal Acquisition Research Symposium. Integrating Theory and Experience: The Acquisition Research Connection*, 1982.

Williams, K. A., and F. Frieman. "Probabilistic Risk Analysis in Conjunction with FAST-E Parametric Cost Estimating." Oak Ridge National Laboratory, CONF-8905119-1, 1989.

Williams, R. F., and R. D. Abeyta (eds). "Management of Risk and Uncertainty in Systems Acquisition: Proceedings of the 1983 Defense Risk and Uncertainty Workshop." Defense Systems Management College, July 1983.

Worm, George H. "Application of Risk Analysis in the Acquisition of Major Systems." Clemson University, 1980.

Worm, George H. "Applied Risk Analysis with Dependence Among Cost Components." In *1982 Federal Acquisition Research Symposium. Integrating Theory and Experience: The Acquisition Research Connection*, 1982.

Young, Philip H. "FRISK — Formal Risk Assessment of System Cost Estimates." American Institute of Aeronautics and Astronautics, AIAA 92-1054, February 1992.

Young, Philip H., and Oliver F. Blackshire "Can the Deadline Be Met? — A Quantification of Schedule Confidence." The Aerospace Corporation, 1986.

ABBREVIATIONS

ABBREVIATIONS

c.d.f.	cumulative distribution function
CER	cost-estimating relationship
CLM	Central Limit Theorem
DoD	Department of Defense
IDA	Institute for Defense Analyses
UHF	ultra-high frequency
USCM	Unmanned Space Vehicle Cost Model
VERT	Venture Evaluation and Review Technique
WBS	work breakdown structure

END

FILMED

DATE: 8-93

DTIC